The Bell System Technical Journal

Vol. XX October, 1941 No. 4

The Reliability of Holding Time Measurements

By ROGER I. WILKINSON

I—The Problem to be Solved

ONE of the fundamental quantities in traffic engineering is the average duration of subscribers' calls. This figure in seconds multiplied by the average number of calls expected over a given route in an hour, and divided by 3600, gives the traffic load submitted in average simultaneous calls—or "the average" as it is commonly called. Tables and curves are widely available which may then be consulted to find the number of paths to be provided so that no more than a desired small percentage of the calls presented will find all paths busy.

The direct measurement of call lengths with a stop watch occurs to one as being the simplest means for obtaining a sample of holding times. It is seldom used, however, due to the relative slowness with which a large number of observations are accumulated coupled with the not inconsiderable expense of the small army of observers required, each looking at one call at a time.

A second direct method of obtaining holding time measurements is by recording mechanically or electrically the length of each call passing over a group of switches or trunks during a certain interval of time. Various holding time recorders or "cabinets" following this principle have been used more or less extensively in the Bell System. Their chief disadvantage has lain in requiring considerable time and labor for summarizing the results. Problems of the perfect maintenance of the measuring equipment have also been present.

To make possible the rapid accumulation of holding time data on a considerable number of calls at relatively slight expense, the method of switch or plug counts has been introduced. This consists in scanning mechanically, electrically, photographically or by eye the group of paths at regular intervals, and recording each time the number found busy.[1] Such data give estimates immediately of the average load being carried, and by a

[1] This number will be highly variable, and even on properly engineered groups great concern need not be felt should few, or even no, cases of "all paths busy" appear since such peaks are of short duration and might easily be missed except in a very long series of counts.

relatively simple analysis a measure of the reliability of such an estimate can be obtained. If in addition for the same period a record is kept on a call- or peg-count meter of the number of calls passing over the group, it is possible also to obtain estimates of the average call holding time and the reliability of such an average.

Direct measurement of holding times or switch counts should naturally be made on groups during periods which are presumably typical of those toward which the engineering is ultimately directed. Usually, although not always, this will be the busy or busiest hours of the day during the busy season of the year. In order to decide intelligently how long a period needs to be studied in any given case some knowledge of the persistence of the same holding time universe is necessary. This might be obtained through relatively small holding time samples made in the hours of interest every day for several weeks in the busy season. If spottiness or "lack of control" is not apparent, the problem will be comparatively simplified. If trends are present, however, it will be necessary to investigate their nature (such as whether some one day of the week shows high holding times) and apportion the main sampling procedure in a fashion to give these peculiarities their proper weighting.

It will be of interest to examine in this respect certain limited data at hand taken by the pen register method some years ago on an inter-office trunk group in Newark, New Jersey. The kind of examination made here will serve to indicate the procedure which may be found suitable in some degree for application to other groups whose characteristics are relatively little known.

II—Preliminary Study of Newark Data

It has long been known that local subscriber call holding times, t, follow remarkably closely the simple exponential frequency distribution,

$$f(t)dt = ke^{-kt}dt, \tag{1}$$

where $\frac{1}{k}$ = the average holding time.[2] This was found to be substantially true of the data collected on the inter-office trunk group in Newark as shown in Fig. 1 for 7385 calls observed in 19 hours having loads in the range 15.0–16.0 average simultaneous calls. The fit of the exponential curve having an average equal to the observed average of 2.380 minutes is seen to be quite good. It may be further noted that in the exponential distribution the standard deviation, σ, equals the mean $\bar{t}$. In practice σ

[2] A. K. Erlang apparently was the first to notice this holding time distribution, "Nyt Tidsskrift for Matematik" (Denmark), 1909.

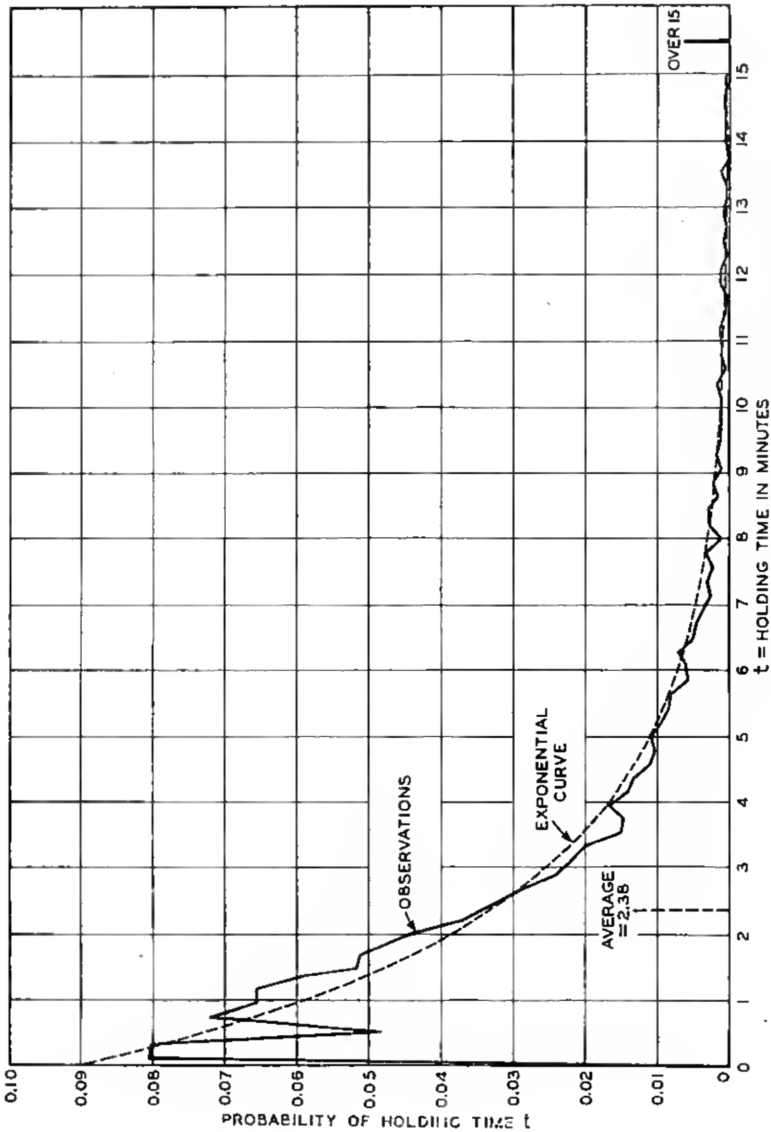

Fig. 1—Distribution of 7385 local holding times, Newark

is usually found to be slightly larger than $\bar{t}$ although not markedly so. We may use this information to test the homogeneity of the holding time universe should all hours of the days be grouped indiscriminately.

TABLE I
HOURLY HOLDING TIME DATA, NEWARK
(Figures in body of table are average holding times in seconds)

Day	Date	Hour of day							
		9-10 am	10-11 am	11-12 am	12-1 pm	1-2 pm	2-3 pm	3-4 pm	4-5 pm
Monday	7- 8-18	131.0	139.0	140.5					
Tuesday	7- 9-18	151.1	151.0	159.0					
Wednesday	7-10-18	146.3	161.4	140.0					
Thursday	7-11-18	138.5	133.7	151.4					
Friday	7-12-18	—	146.0	138.1					
Saturday	7-13-18	123.7	139.5	130.5					
M.	7-15-18	135.1	152.6	147.0					
T.	7-16-18	134.5	138.4	—					
W.	7-17-18	138.4	151.3	159.8					
Th.	7-18-18	148.2	148.0	—					
F.	7-19-18	147.1	136.4	—					
S.	7-20-18	146.9	131.8	—					
M.	7-22-18	145.2	146.7	148.7					
T.	7-23-18	154.5	145.3	143.7					
W.	7-24-18	132.5	137.7	157.5					
Th.	7-25-18	—	149.0	—					
F.	7-26-18	138.0	157.9	174.4					
S.	7-27-18	128.5	142.0	150.1					
M.	7-29-18	132.3	141.5	—		166.7			
T.	7-30-18	151.3	143.4	139.5					
W.	7-31-18	142.1	129.4	144.1					
Th.	8- 1-18	141.1	134.4	154.1					
F.	8- 2-18	161.6	150.5	150.3					
S.	8- 3-18	148.7	147.7	134.5					
M.	8- 5-18	139.4	131.0	142.9					
T.	8- 6-18	158.0	141.4	158.5					
T.	8-13-18	162.8	141.6	—					
W.	8-14-18	136.0	150.8	—					
Th.	8-15-18	153.0	141.8	139.0					
F.	8-16-18	141.0	160.1	151.6					
Th.	9- 5-18	144.7	158.9	—					
F.	9- 6-18	—	139.5	—					
W.	9-25-18	—	—	144.3				157.6	
Th.	9-26-18	152.1	143.1	134.2					
F.	9-27-18	139.7	160.5	149.9					
M.	9-30-18	—	132.8	128.9		138.7			152.6
T.	10- 1-18	129.7	137.5	150.0		158.3			166.0
W.	10- 2-18	138.5	135.2	132.5			161.0		
Th.	10- 3-18	142.0	143.0	152.0	174.3	153.0	165.0		
F.	10- 4-18	128.4	136.9	145.7		150.5			
S.	10- 5-18	137.0	138.4	138.5					
M.	10- 7-18	—	—	—		150.0	139.2	137.7	
T.	10- 8-18	131.0	136.4	145.1			150.0	145.1	152.2
W.	10- 9-18	138.3	144.4	142.0		174.8		145.4	151.6
Th.	10-11-18	135.4	149.8	—					
Summary:									
No. Hours		39	43	33	1	7	4	4	4
Average		141.63	143.67	146.01	174.3	156.3	153.8	146.4	155.6

In Table I and on Fig. 2 are shown the holding time averages for 135 hours observed at various times of the day over a period of 3 months. At first glance these appear to fall in two rather distinct groups, those before noon and those after noon. If the 115 hours before noon be considered as defining a homogeneous group, could those holding time averages found in the afternoon be reasonably considered as coming from the same universe? We first find the average holding time of the 115 forenoon hours to be 143.5 seconds. Since these hours averaged about $n = 390$ calls each, the standard deviation of the means $\sigma_{\bar{t}}$ should, by theory, be closely

$$\sigma_{\bar{t}} = \frac{\sigma}{\sqrt{n}} = \frac{\bar{t}}{\sqrt{390}} = 7.29.$$

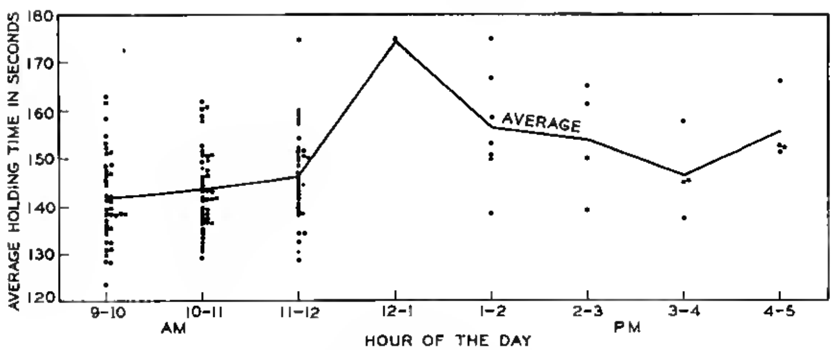

Fig. 2—Day to day holding time averages by hours of the day, 135 hours, Newark

The standard deviation observed is 9.26, some 27% higher, which, however, agrees with the observation made in the previous paragraph. On the hypothesis that the universe of 115 early hours has the parameters of $\bar{t} =$ 143.5 and $\sigma = 7.29$, we see that the observations for each of these three clock hours could readily have occurred. The deviations of their averages from 143.5 are 1.9, .17 and 2.5 seconds, respectively, and according to theory the corresponding standard errors in these averages are 1.168 $\left(= \frac{7.29}{\sqrt{39}}\right)$, 1.111 $\left(= \frac{7.29}{\sqrt{43}}\right)$, and 1.270 $\left(= \frac{7.29}{\sqrt{33}}\right)$. All the deviations are well within two times the standard error of the assumed mean of the holding time universe. The remaining 20 observations from noon on, however, average 154.6 seconds, and if they could reasonably have come from the hypothesized universe, this figure should not differ from 143.5 by more than, say, three times the standard error 1.630 $\left(= \frac{7.29}{\sqrt{20}}\right)$. Actually the difference is more

than six times the standard error, strongly indicating a significant difference between the forenoon and afternoon holding times. We conclude that between 9 a.m. and noon the holding times are satisfactorily controlled but that we should not attempt to include observations on afternoon hours with them. Since the heaviest loads here occurred generally in the morning we should confine our direct measurements or switch counts to these hours for determining the engineering holding time.

It may occasionally be well to investigate the possibility that certain days of the week have, on the average, longer holding times than other days. If the Newark 9–12 a.m. data are plotted by days of the week as in Fig. 3 we see that the averages for each day fluctuate considerably as shown by the heavy dots. In testing these points the simple average of the σ's for

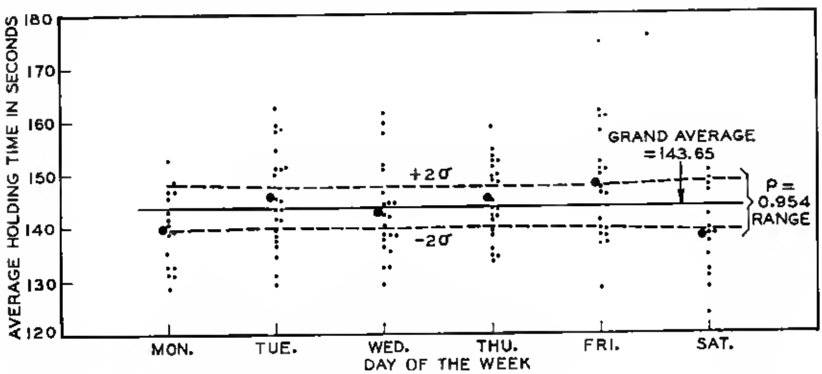

Fig. 3—Variations in average holding times, 115 9–12 a.m. hours, by days of the week, Newark

each day's hours is taken as an estimate of the standard deviation of the homogeneous universe from which all the hours are presumed to be drawn. Then with the weighted arithmetic means of the daily averages as the best estimate available for the mean of the universe, each day's average is tested to see whether it could reasonably have arisen from it. The $\pm 2\sigma$ lines which should include some 95% of the day-averages are shown on Fig. 3. It is seen that two of the six points fall slightly outside these limits indicating a moderately significant difference in the holding time conditions for Friday and Saturday. The sampling procedure to follow in such cases of non-controlled populations is not rigorously definable. However, it is clear that the samples should be drawn from the various groups of controlled elements which probably go to make up the universe, and rougbly in proportion to the importance to be assigned to each such group. In our example here we would probably want to draw samples of about equal size from the calls of each week day in the week.

Finally there may be some question as to the busy season, its length and stability. Plotting of the same data for the 9–12 a.m. hours as in Fig. 4A and 4B will help to decide these points. The morning hours' holding time averages for each day of the week are plotted for several weeks during the suspected busy season. Wide changes in the load through the passage

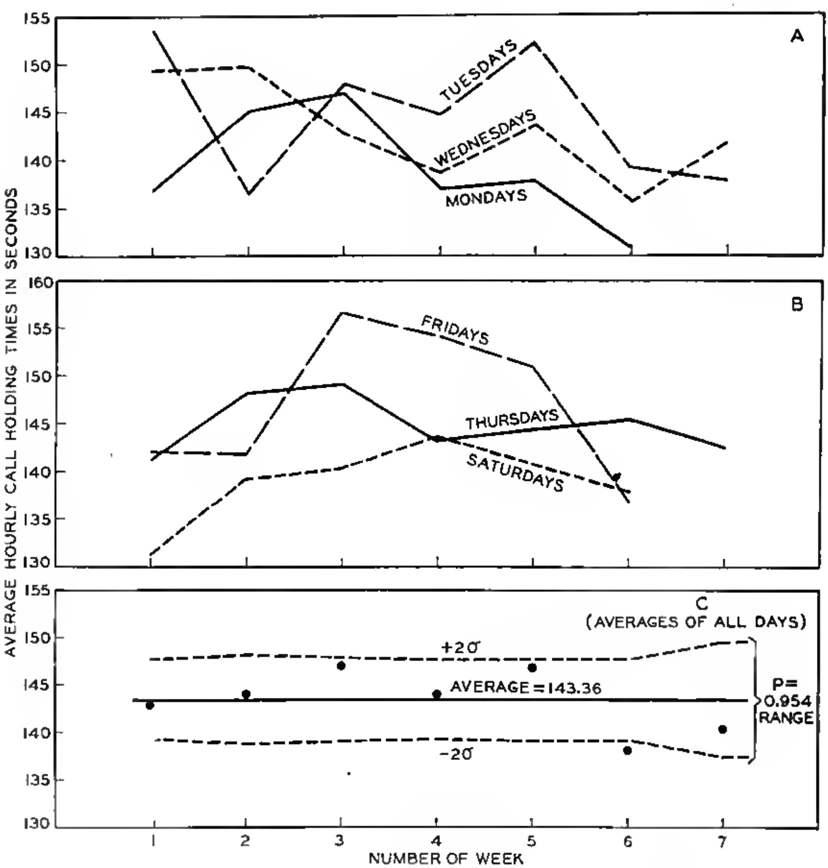

Fig. 4—Seasonal trend in holding time averages, 105 9–12 a.m. hours, Newark

of weeks can be noted by eye. In the case illustrated there appears no consistency of movement. Applying the identical test used in the previous paragraph for the day-of-the-week changes, we find in Fig. 4C that in the first five-week period no movements of significance took place. However, the sixth week, which followed the first five after an interval of about six weeks, showed a significant drop suggesting the approach of a lower

level of traffic. The traffic engineer would probably decide to schedule his holding time observations during the weeks numbered one to five inclusive.

Having determined something as to the character of the holding time trends, if any, with hours of the day, with days of the week, and seasons of the year the traffic engineer is in a better position to lay out a program for sampling. He will especially want to apportion the total sample between the hours or days which show significant differences among themselves roughly in proportion to the relative traffics flowing at those levels. The less specific the information on the traffic flow characteristics the more important it will be to spread the observations over a variety of hours, days or weeks.

III—A Satisfactory Sample of Directly Measured Holding Times

If the standard deviation σ of individual call lengths is known, we can estimate the standard error of the average of n measurements as

$$\sigma_{\text{avg}} = \frac{\sigma}{\sqrt{n}}. \tag{2}$$

Since n will usually be several hundred we can obtain a good figure for σ by calculating the standard deviation, S, of the n observations. As noted before, for exponential calls this will be not far from the average holding time l which may be substituted for σ if great accuracy is not required. In fact if the sampling is representatively made from a universe not strictly homogeneous, the better figure for σ may be the average l, instead of the standard deviation found in the sample since in so-called Poisson Sampling of stable but nonhomogeneous universes the standard error of the average may be somewhat reduced from $S/\sqrt{n}$.

We may now make the statement that for n large the probability is P that the true average holding time does not differ from that observed by more than $\pm z \dfrac{l}{\sqrt{n}}$ seconds, where P and z are given in the table below.

TABLE II

P	z	P	z
.50	.6745	.95	1.960
.85	1.440	.99	2.576
.90	1.645	.999	3.291

For example if we have measured the individual lengths of 900 calls which show an average of 150.3 seconds, we are then 99% sure that the true holding time average for the sampled universe lies closely within the range

$$150.3 \pm 2.576 \frac{150.3}{\sqrt{900}},$$

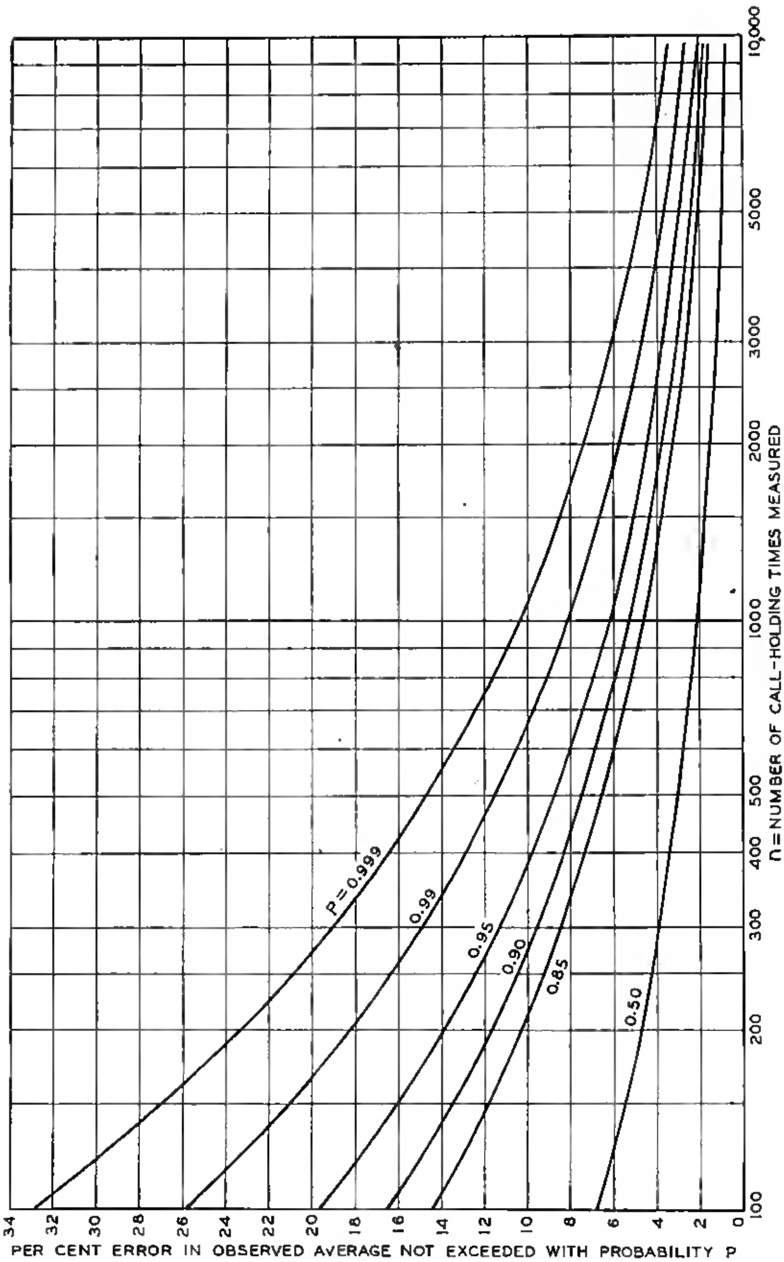

Fig. 5—Error in exponential-holding-time average determined by stop watch method

that is between 137.4 and 163.2 seconds. (The best single estimate, of course, is the observed average of 150.3 seconds.) Or conversely, if one should desire to determine the true average holding time within 5 seconds with a surety of $P = .90$, he may use an approximate value of the standard deviation (or the average holding time) based on past experience and substitute in

$$n = \left(\frac{1.645\bar{l}}{5}\right)^2,$$

to obtain the number of calls to be measured. If $\bar{l} = 150$ seconds here, $n = 2421$.

The same information is contained in Fig. 5 which gives the *per cent* error, $\pm 100 z_P/\sqrt{n}$, in the observed average not exceeded with probability P.

All this is based on the assumption that each of the n call lengths is accurately enough measured so that no appreciable error is introduced from this source. Obviously there is no point in expending much effort in carefully "proportioning" a sample so as to be representative of the vagaries of the universe if each of the calls so chosen is not pretty accurately measured. This would be quite as futile as measuring very accurately the holding times of a number of calls chosen during some short time period which might turn out to be wholly untypical of certain of those important periods coming earlier or later. For these direct measurement cases it will probably be quite satisfactory if each call is measured with a maximum error of not over one-tenth of $\bar{l}/\sqrt{n}$. In our example of 900 calls this would be .501 seconds, that is measurement of each call to the nearest second.

IV—Holding Times by Switch Count Methods

If each call's holding time is not measured with considerable accuracy it is immediately clear that additional calls must be observed in order to compensate therefor. This is the situation in the method of switch counts which is in effect a means for noting at regular intervals i whether a particular call does or does not exist. Thus none of the calls are at all closely measured for their individual lengths. Other errors will also have to be considered since at the beginning of the period some switch counts are inevitably included on a number of calls from the preceding hour and at the end of the period some of the calls registered on the peg count meter will end beyond the period with the loss of part of their proper switch counts. As a result there are in this method three distinct sources of holding time error whose magnitudes we shall proceed to investigate in turn:

a. Errors at the start of the observation period;
b. Errors at the end of the observation period;
c. Errors at the beginning and end of each call.

The theoretical conclusions will be compared at various points with certain data available.

a. Errors at the Start of the Observation Period

If the period of observation T be divided into r equal intervals of length i, and switch counts are made at the beginning (and end) of each interval, we shall have a total of $r + 1$ observations. The #1 count will give us immediately the number of calls extending into the period from the pre-

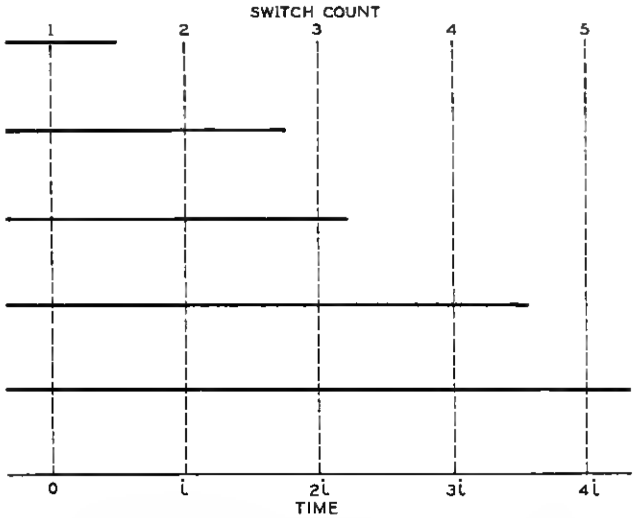

Fig. 6—Diagram of switch counts at beginning of the period

ceding hours. We have no means, however, of segregating their contributions to subsequent switch counts, so a theoretical estimate of this amount is required.

If the average holding time of the calls is $\bar{t}$, and they follow closely the exponential law of distribution, we may reason as follows.[3] Consider the case of a single call passing time 0 at the start of the observation period as in Fig. 6. Then the probability that it will be included only in switch count number 1, that is that it ends between time 0 and time i, is

$$p_1 = P(<i) = 1 - e^{-\frac{i}{\bar{t}}}.$$

[3] Of course we do not know $\bar{t}$ exactly since that is the ultimate object of our study; however, for the present purpose great accuracy will not be required, and $\bar{t}$ can usually be taken as the first estimate of holding time obtained by the switch count method without corrections.

Similarly the probability that exactly two switch counts will be contributed by such a call is

$$p_2 = P(>i) - P(>2i) = e^{-\frac{i}{t}} - e^{-\frac{2i}{t}}.$$

Likewise,

$$p_3 = P(>2i) - P(>3i) = e^{-\frac{2i}{t}} - e^{-\frac{3i}{t}},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$p_u = P[>(u-1)i] - P(>ui) = e^{-\frac{(u-1)i}{t}} - e^{-\frac{ui}{t}}.$$

If now there have been m such calls observed on switch count number 1, we shall need to add m variables of the type

$$f(u) = e^{-\frac{(u-1)i}{t}} - e^{-\frac{ui}{t}} = e^{-\frac{ui}{t}}(e^{\frac{i}{t}} - 1) = ce^{-\frac{ui}{t}},$$

where u may take all values from 1 to $r + 1$. The exact addition of these variables when m is more than a small number, say 3 or 4, becomes quite complex. However, in such cases (which may be the rule) we revert to the method of combining their individual moments to obtain the moments (and parameters) of the resultant distribution. We find for a single variable,

usage # of S.C. calls carried into measured period

$$\bar{u} = c\sum_{u=1}^{r+1} ue^{-u\frac{i}{t}} = \frac{1}{1 - e^{-\frac{i}{t}}}\left[1 - e^{-(r+1)\frac{i}{t}}\right], \tag{3}$$

and

$$\sigma_u = \frac{e^{-\frac{i}{2t}}}{1 - e^{-\frac{i}{t}}}\left[1 - 2(r+1)\left(1 - e^{-\frac{i}{t}}\right)e^{-r\frac{i}{t}} - e^{-(2r+1)\frac{i}{t}}\right]^{\frac{1}{2}}. \tag{4}$$

The factors shown in brackets in equations (3) and (4) will approach unity very closely in any practical applications of the present type; they will therefore be omitted in the subsequent analysis.

The mean and standard deviation of the sum of m such variables are readily determined, of course, as[4]

[4] It is interesting to note that if the first switch count had been omitted so that only $\bar{m}$ could have been estimated from the average of the switch counts from #2 onward, we might have assumed a Poisson distribution for m, that is $p_m = \frac{\bar{m}^m e^{-\bar{m}}}{m!}$, and thereby have obtained an estimate of the switch counts contributed by calls from the preceding period as follows,

$$\bar{s}'_m \approx \frac{\bar{m}}{1 - e^{-\frac{i}{t}}}, \tag{5'}$$

$$\sigma'_m \approx \frac{\sqrt{\bar{m}(e^{-\frac{i}{t}} + 1)}}{1 - e^{-\frac{i}{t}}}. \tag{6'}$$

$$\hat{s}_m = m\bar{u} = \frac{m}{1 - e}, \tag{5}$$

$$\sigma_m = \sqrt{m}\,\sigma_u = \sqrt{m}\,\frac{e^{-\frac{i}{2t}}}{1 - e^{-\frac{i}{t}}}. \tag{6}$$

In Fig. 7 is shown a comparison between this theory and the actual numbers of 1-minute switch counts contributed by these carry-over calls

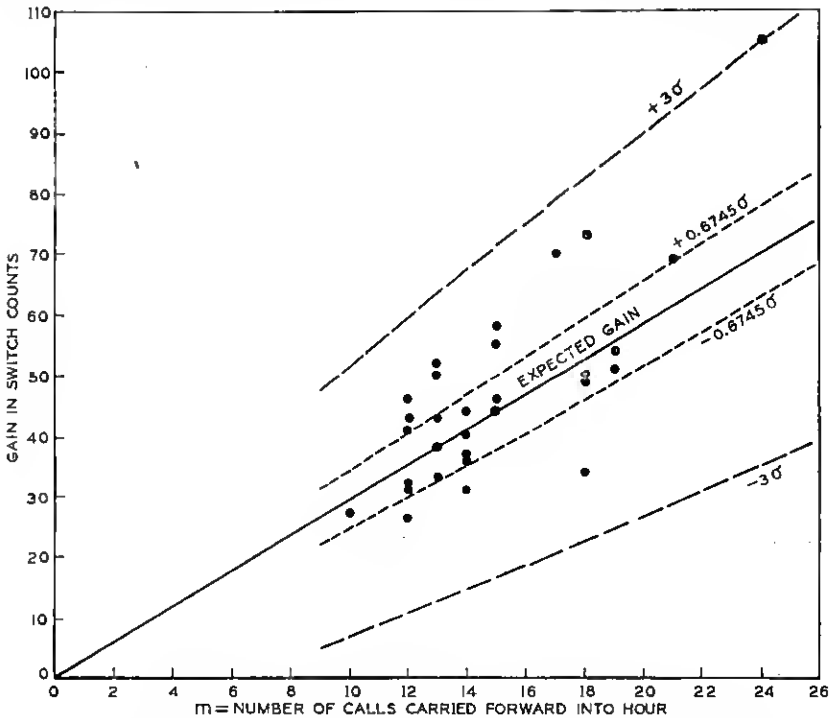

Fig. 7—Gains over true switch counts at the start of the observation period

in 30 hourly periods of observation on the Newark interoffice trunk group previously described. Since $\bar{m}$, the average number carried forward is about 15, it was thought that the sum-distribution of so many variables would rather closely approach normality. On this assumption the ±50% and ±99.73% "control" lines have been drawn on Fig. 7. One point falls on the latter limit lines, while 16 of the 30 fall within the 50% lines, thus providing a gratifying corroboration of the theory.

b. Errors at the End of the Observation Period

If switch counts have been made at regular intervals i so that the $r + 1$st count occurs at the exact end of the period for which the number of originating calls has been registered, then the situation closely resembles that at the start of the period. The $r + 1$st observation tells us immediately how many calls are continuing into the next hour. A particular one of these calls may extend to the areas 0 to i, i to $2i$, $2i$ to $3i$, etc., measured beyond the end of the period as in Fig. 8. A call ending in the interval $2i$ to $3i$,

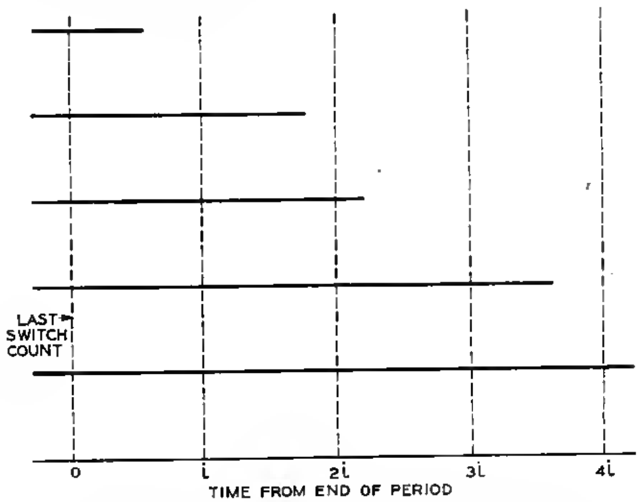

Fig. 8—Diagram of switch counts at the end of the period

for instance, would fail to have two switch counts marked up if the counting stopped at 0. Then the probability of losing exactly zero switch counts is

$$p_0 = P(<i) = P(>0) - P(>1i) = 1 - e^{-\frac{i}{\bar{i}}},$$

and in general

$$p_v = P(>vi) - P[>(v+1)i] = e^{-\frac{vi}{\bar{i}}} - e^{-\frac{(v+1)i}{\bar{i}}} = e^{-\frac{vi}{\bar{i}}}\left(1 - e^{-\frac{i}{\bar{i}}}\right)$$

$$= c'e^{-v\frac{i}{\bar{i}}} \tag{7}$$

where v varies from o to ∞. The average and standard deviation of a single variable will be[5]

$$\bar{v} = \sum_{v=0}^{\infty} v p_v = \frac{e^{-\frac{i}{\bar{i}}}}{1 - e^{-\frac{i}{\bar{i}}}}, \tag{8}$$

[5] The value $\bar{v}$ is one less than $\bar{u}$ shown in equation (3) after neglecting the minute correction factor, since each call there received a switch count at 0 time which is omitted here. The standard deviation is identical with equation (4) without the correction factor, since a constant deduction has simply been made on each call.

and

$$\sigma_v = \frac{e^{-\frac{i}{2\bar{i}}}}{1 - e^{-\frac{i}{\bar{i}}}}. \tag{9}$$

The frequency distribution of the sum of w such discrete variables v is readily found to be

$$f(v_w) = \left(1 - e^{-\frac{i}{\bar{i}}}\right)^w \frac{(w + v_w - 1)!}{(w-1)!\, v_w!} e^{-v_w \frac{i}{\bar{i}}}.$$

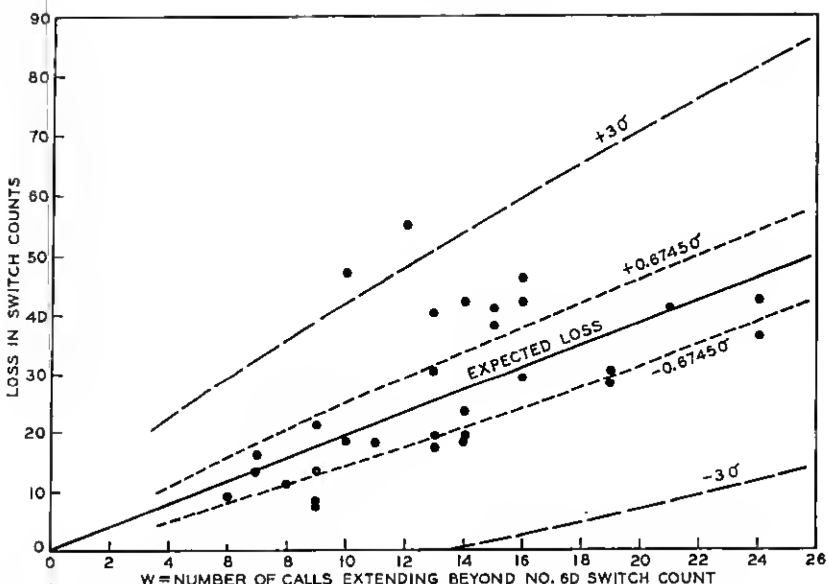

Fig. 9—Loss of true switch counts at the end of the observation period

The parameters of this sum distribution are

$$\bar{s}_w = w\bar{v} = \frac{we^{-\frac{i}{\bar{i}}}}{1 - e^{-\frac{i}{\bar{i}}}}, \tag{10}$$

$$\sigma_w = \sqrt{w}\,\sigma_v = \sqrt{w}\,\frac{e^{-\frac{i}{2\bar{i}}}}{1 - e^{-\frac{i}{\bar{i}}}}. \tag{11}$$

A check on these last formulas is shown in Fig. 9 where the numbers of switch counts lost on w calls carried beyond the last (or r + 1st) switch count are recorded for 30 busy hour periods. Two of the hours show results outside the theoretical $\pm 3\ \sigma$ normal curve limits but the falling of almost

exactly half of the points within the ±50% lines reassures us that the estimated parameters are probably quite good. (If instead we had calculated, say, ±99% limit lines based on the skew distribution $f(v_w)$ just derived, it seems likely that even the two "unusual" points of Fig. 9 might have fallen in the "reasonable" range.)

At this point it will be well to point out that in making switch counts considerable care needs to be exercised in two directions. First, switch counts #1 and #$r + 1$ should coincide very closely with the beginning and end, respectively, of the observation period, the intermediate counts of course being uniformly spaced. Second, each count should be taken as quickly as possible so that a substantially instantaneous reading of calls in progress is obtained. These two desiderata can usually be attained readily in schemes using mechanical, electrical or photographic means for recording the switch counts. Counts made by observers, however, may be subject to highly variable errors since in some cases a substantial portion of the interval i may be required to complete a count. Such uncertainties naturally increase the end-effect errors, and, consequently, the overall error in an average-holding-time determination.

To estimate the magnitude of the increased errors resulting from failure to meet the above switch count specifications would require a special study for each kind and type of failure; these would probably differ in every application. An idea of the sensitiveness of switch counting to such irregularities and the likely order of magnitude of the increased errors may be gained by examining certain of the Newark data. Here the last switch count in many of the hours, although taken instantly, failed to coincide well with the end of the observation period T. The last count fell at points varying from a little after the period closed to nearly an interval i ahead of this instant as shown in Fig. 10. In most hours this permitted a small number of calls to mark up on the peg count register after the last switch count was taken.

As a result, if corrections were not made, the estimate of switch counts lost on calls extending beyond the end of the hour would have omitted about a third of those properly included, with a consequent lowering of the average holding time estimate by approximately one per cent.

If the time j which has elapsed between the last switch count and the end of the period T is known, certain corrections for this particular irregularity can be attempted. We shall indicate the formulas required since they will be useful in an analysis of the Newark data. If an average of $\alpha = \frac{j}{i} a$ calls are assumed to originate in the interval j, and they follow a Poisson distribution $\frac{\alpha^x e^{-\alpha}}{\underline{/x}}$, then the expected number of switch counts lost

is found to be

$$E = \alpha \frac{\frac{\bar{t}}{j}\left(e^{\frac{i}{\bar{t}}} - 1\right)\left(1 - e^{-\frac{j}{\bar{t}}}\right)e^{\frac{j-2i}{\bar{t}}}}{\left(1 - e^{-\frac{i}{\bar{t}}}\right)^2}, \tag{12}$$

and the standard deviation is

$$\sigma' = \sqrt{\alpha}\left[\frac{\bar{t}}{j}\left(e^{\frac{i}{\bar{t}}} - 1\right)\left(1 - e^{-\frac{j}{\bar{t}}}\right)e^{\frac{j-2i}{\bar{t}}}\,\frac{1 + e^{-\frac{i}{\bar{t}}}}{\left(1 - e^{-\frac{i}{\bar{t}}}\right)^3}\right]^{\frac{1}{2}}. \tag{13}$$

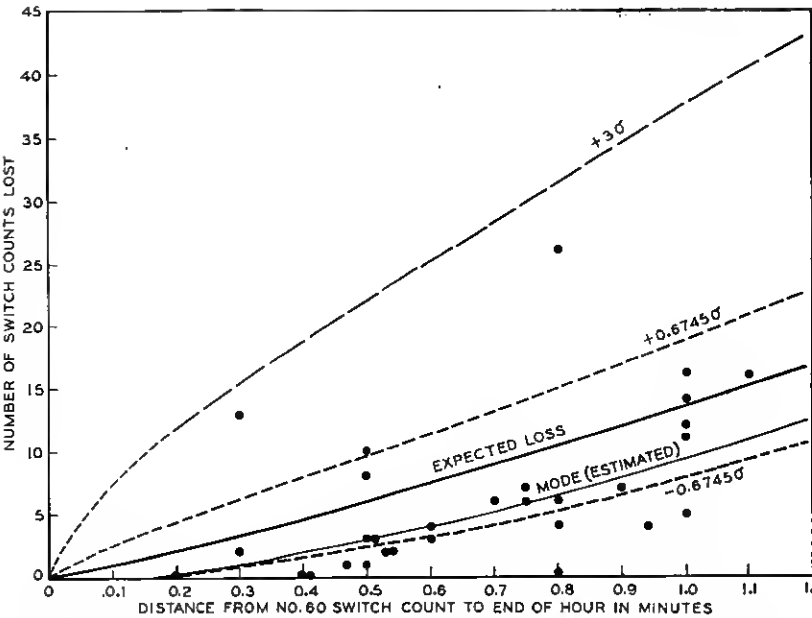

Fig. 10—Switch counts lost on calls which originated after the last switch count

If j is small, α will be correspondingly small making the average switch counts lost, E, small. The distribution of lost switch counts will then be very skew since the case of zero lost will be prominent. For larger j's the distribution should assume a unimodal form, gradually becoming less skew. A number of Newark busy hours were studied in this fashion, and the results are compared with theory on Fig. 10. The agreement is seen to be fair and a modal line estimated by eye falls substantially below the expected mean corroborating the decided skewness just predicted. The wide fluctuations in lost switch counts, each one of which if incorrectly estimated results in a considerable error (in the present example about .2 second)

in the estimate of the average holding time, will serve to indicate the desirability where possible of eliminating altogether this and other supplementary errors by seeing that individual counts are taken very quickly, and that the first and last switch counts coincide closely with the ends of the observation period.

c. Errors in Measuring Each Call

Due to the method of counting the switches at finite intervals, an exact measurement of the length t of any one call will seldom if ever be made. We shall attempt in what follows to determine the magnitude and char-

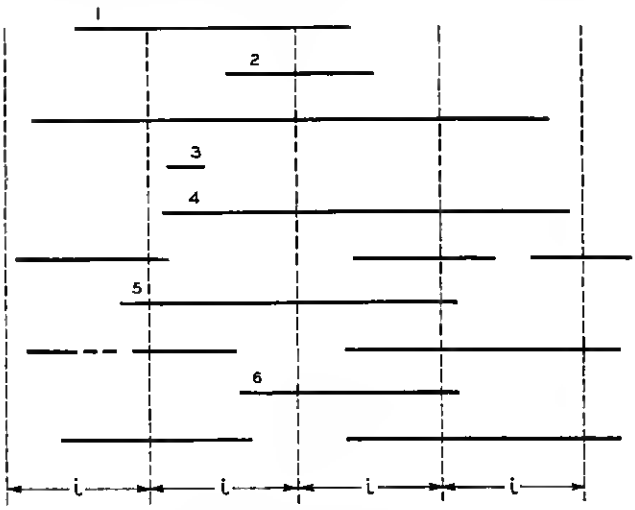

Fig. 11—Typical field of switch counts

acteristics of these errors in measuring individual calls. In any field of switch counts, such as in Fig. 11, there will be calls of types #1 and #2 which receive about one switch count for each i call seconds of length. There will likewise be many others such as #3 and #4 which will be substantially undercounted, and about as many others, #5 and #6, which will be overcounted. We shall proceed with certain special cases, and then obtain the general result desired.

Case 1. t *lies between* 0 *and* i

If the holding times t follow some law of fluctuation $f(t)$, a certain proportion of them will have lengths lying in the range $t = 0$ to $t = i$. Such a call will either cross one of the switch count points, or it will not. Upon the assumption of a random instant of origination the probability of its

crossing will be simply $\frac{t}{i}$. That is, if the origination point falls within t seconds to the left of a switch count, the call will be marked up; if not it will not be counted. If it is marked up there will occur for that call a plus estimation error of $x = i - t$ seconds since we shall eventually assume that every switch count infers i call seconds of use on the trunk group. Likewise there is a probability of $\frac{i - t}{i}$ that the call will not be counted with the resultant negative error of $x = -t$ seconds. In summary the total probability of a positive error of $x = i - t$ is

$$f(t)\ dt\ \frac{t}{i};$$

and for a negative error of $x = -t$,

$$f(t)\ dt\ \frac{i - t}{i}.$$

Case 2. t lies between i and $2i$

Such a call may be included either once or twice in the switch counts. The probability that it is counted twice is $\frac{t - i}{i}$ with a resultant plus error of $x = 2i - t$. The probability that it is counted but once is $1 - \frac{t - i}{i} = \frac{2i - t}{i}$, with the corresponding negative error of $x = i - t$. The overall probabilities of course will be formed by weighting these as in the first case with the probability, $f(t)dt$, that the holding time of length t to $t + dt$ actually occurs.

General Case. t lies between qi and $(q + 1)i$

It will readily be seen that by extending the reasoning of the two cases above to the case of t lying between qi and $(q + 1)i$ we shall have a plus error (due to the call being marked up $q + 1$ times) of $x = (q + 1)i - t$, with a probability of occurrence of $\frac{t - qi}{i}$, and a negative error (the call marked up q times) of $x = qi - t$ with a probability of $\frac{(q + 1)i - t}{i}$.

Summarizing the above cases, a *negative* error of size x can occur in a great number of ways, due to t taking the values $-x$, $i - x$, $2i - x$, $\cdots$ $qi - x$, $\cdots$ with the corresponding probabilities of occurrence of the call lengths, $f(-x)$, $f(i - x)$, $f(2i - x)$, $\cdots$ $f(qi - x)$ $\cdots$, respectively. In

addition each time such a call length does occur we must introduce the contingent probability $\frac{i+x}{i}$ that a negative and not a positive error will occur. The total probability of making an error of x, where $x \leqq 0$, on any call is then,

$$p_{x \leqq 0}(x)\,dx = \frac{i+x}{i}\,[f(-x) + f(i-x) + f(2i-x) + \cdots]\,dx. \quad (14)$$

Similarly we find the total probability of making a positive error of magnitude x, on any call, as

$$p_{x \geqq 0}(x)\,dx = \frac{i-x}{i}\,[f(i-x) + f(2i-x) + f(3i-x) + \cdots]\,dx. \quad (15)$$

It will now be of interest to apply equations (14) and (15) to some particular types of holding time distributions.

(*a*). *Constant Holding Times*, $t = h$[6]

If t is constant and equal to h, it will necessarily fall within some one of the special cases enumerated above. Suppose h lies between qi and $(q+1)i$. There will be but one value of the error x_1 possible in the negative range and it will equal $qi - h$, with a corresponding single value x_2 in the positive range equal to $(q+1)i - h$. It will be seen that equations (14) and (15) reduce simply to

$$p_{x_1 \leqq 0}(x_1) = p(qi - h) = \frac{i+x_1}{i}\,f(qi - x) = \frac{i+x_1}{i}\,f(h) = \frac{i+x_1}{i}, \quad (16)$$

and

$$\begin{aligned} p_{x_2 \geqq 0}(x_2) &= p[(q+1)i - h] \\ &= \frac{i-x_2}{i}\,f[(q+1)i - x_2] = \frac{i-x_2}{i}\,f(h) = \frac{i-x_2}{i}. \end{aligned} \quad (17)$$

The mean and standard deviation of this two-valued variable are found to be

$$\bar{x} = 0, \quad (18)$$

$$\sigma_x = \sqrt{-x_1(i+x_1)} = \sqrt{(i-x_2)x_2} = \sqrt{-x_1x_2}. \quad (19)$$

It may be noted that σ_x attains a maximum of $i/2$ when $x_2 = i/2$, and approaches 0 for $x = 0$. This is of importance when one has to choose an observation interval i for switch counting constant or relatively constant holding times.

[6] The particular error distributions for cases *a* and *c* were obtained by G. W. Kenrick in 1923.

Example: An hour with 372 calls having a constant holding time per call of $h = 131.8$ seconds was subjected to a 60 second switch count study, records being kept of the errors in measurement on individual calls. As shown in Fig. 12, 284 or 76.3% gave counts of "2" with an error of $120 - 131.8 = -11.8$ seconds. The remaining 88 calls, or 23.7% received counts

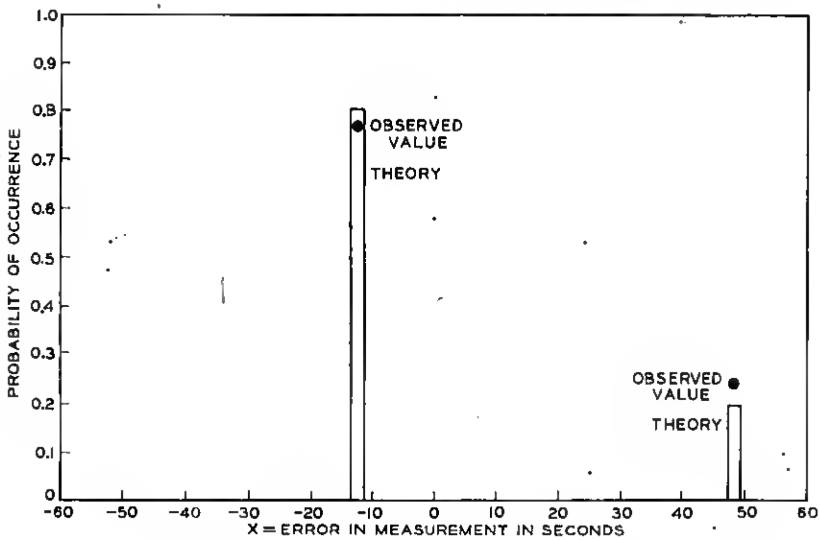

Fig. 12—Error distribution for measurements on individual calls with constant holding times

of "3", with errors of $180 - 131.8 = 48.2$ seconds. Applying the theory just developed to this case gives,

$$p(2i - h) = p(-11.8) = \frac{60 - 11.8}{60} = .803,$$

$$p(3i - h) = p(48.2) = \frac{60 - 48.2}{60} = .197.$$

As indicated on Fig. 12, these theoretical values check very satisfactorily with the observations. The observed average holding time = 134.2 seconds as against the true value of 131.8 seconds; the error of 2.4 seconds is quite compatible with $\sigma_x = \sqrt{11.8(48.2)} = 23.85$ seconds and the $n = 372$ calls observed.

(b). Equally Likely Distribution of Holding Times between Adjacent Multiples of i

Imagine a holding time distribution of any general form but with a constant probability of occurrence between adjacent pairs of multiples of i,

such as in Fig. 13. Such a distribution would probably occur but rarely, if ever, in practice. However, if the intervals i are short compared to the average holding time $\bar{l}$, such an assumption may not introduce any serious discrepancy in whatever form is simulated.[7]

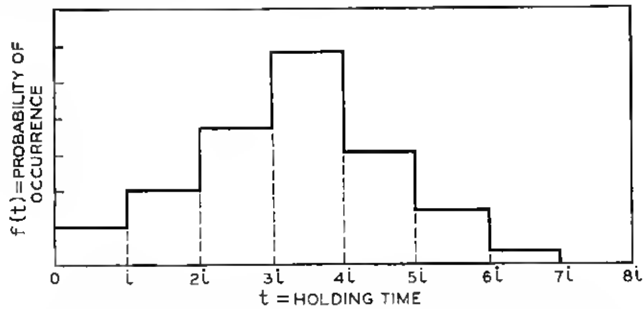

Fig. 13—A varying holding time with a number of equally likely ranges

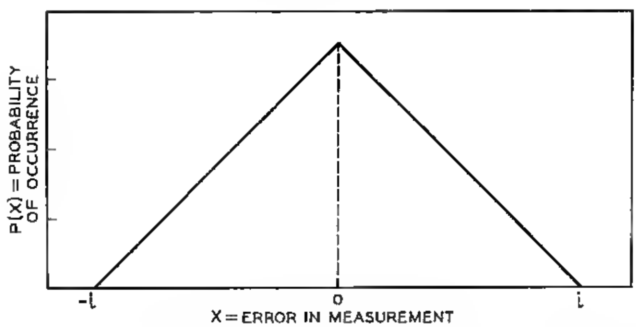

Fig. 14—Distribution of call measurement errors for "equally likely" holding time distributions

In this case it is obvious upon inspection that the sum of the terms in the brackets in equation (14) is a constant for all values of x, and likewise in equation (15), and that they equal each other. Hence

$$p_{x \leq 0}(x) = K \frac{i + x}{i}, \tag{20}$$

and

$$p_{x \geq 0}(x) = K \frac{i - x}{i}, \tag{21}$$

[7] The analytics of the allied case in which the errors at the ends of a call were assumed to fall equally likely between $\pm i/2$, were discussed by E. C. Molina in an unpublished memorandum dated September 7, 1920.

give the isosceles triangular distribution of errors on individual calls shown in Fig. 14. In this the average error is

$$\bar{x} = 0, \tag{22}$$

and the standard deviation is

$$\sigma_x = .408i. \tag{23}$$

(c). Holding Times Exponentially Distributed, $f(t) = ke^{-kt}$, *where* $k = \frac{1}{\bar{t}}$

With holding times of the exponential type the sum of the terms in the brackets of equations (14) and (15) will depend on the particular magnitudes of the errors x assumed. If in equation (14), we substitute exponential expressions for the f-functions, we have

$$\begin{aligned} p_{x \lessgtr 0}(x)\,dx &= \frac{i+x}{i}\left(ke^{-k(-x)} + ke^{-k(i-x)} + ke^{-k(2i-x)} + \cdots\right) dx \\ &= \frac{i+x}{i}\,ke^{kx}\left(1 + e^{-ki} + e^{-2ki} + \cdots\right) dx \\ &= \frac{i+x}{i}\,ke^{kx}\,\frac{1}{1-e^{-ki}}\,dx \\ &= k'\,\frac{i+x}{i}\,e^{\frac{x}{\bar{t}}}\,dx, \end{aligned} \tag{24}$$

where

$$k' = \frac{1}{\bar{t}\left(1 - e^{-\frac{i}{\bar{t}}}\right)}.$$

Similarly we find

$$p_{x \gtrless 0}(x)\,dx = k'\,\frac{i-x}{i}\,e^{-\frac{i-x}{\bar{t}}}\,dx. \tag{25}$$

The mean and standard deviation of this unusual-shaped distribution of x are found to be

$$\bar{x} = 0, \tag{26}$$

$$\sigma_x = \sqrt{\bar{t}}\,\sqrt{\frac{(2\bar{t}+i)e^{-\frac{i}{\bar{t}}} - 2\bar{t} + i}{1 - e^{-\frac{i}{\bar{t}}}}}. \tag{27}$$

In Fig. 15 is shown the distribution of the individual errors found by 60-second switch counts on 746 varying holding time calls (2 hours on the

Newark group). Their true average holding time was 131.45 seconds. The mean error was found to be +1.84 seconds and the standard deviation 25.55 seconds. The corresponding theoretical distribution is found to have a standard deviation of 24.56 seconds with a mean, of course, of zero. The theoretical distribution is superposed on the data of Fig. 15 and is seen to give quite a good fit.

It is interesting that the theoretical average error for each of these three widely dissimilar holding time distributions should be zero, while their standard deviations and analytical forms assume quite different characteris-

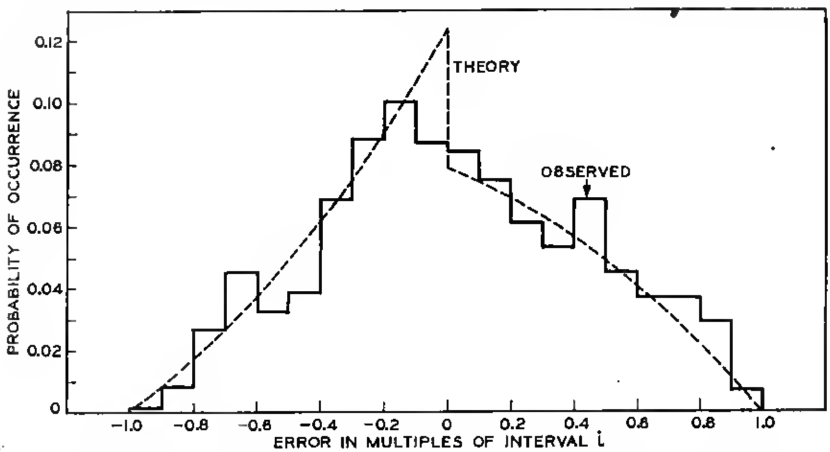

Fig. 15—Distribution of call measurement errors for exponential holding times

tics.[8] A comparison of the σ's obtained from equations (19), (23) and (27) for a typical choice of values, $\bar{l} =$ 145 seconds, $i =$ 60 seconds, gives

$$\sigma_x \text{ constant h.t.} = 29.58 \text{ seconds,}$$

$$\sigma_x \text{ ``Equally likely'' h.t.} = 24.48 \text{ seconds,}$$

$$\sigma_x \text{ Exponential h.t.} = 24.03 \text{ seconds.}$$

The σ_x constant h.t. is largely a function of whether l is closely a whole multiple of i; comparing the values for the other two σ_x's indicates it is slightly advantageous that most of the variable holding time calls to be switch counted in practice are of a roughly exponential form.

The Total Error on n Calls

We shall now attempt to combine the errors from the three sources just discussed and formulate some general conclusions for making the most

[8] It may readily be shown that the average error will be zero for any assumed holding time distribution, by noting that each length of call therein may momentarily be segregated and considered under paragraph "a" as a constant holding time.

of the switch count method of estimating average holding times. Suppose we have completed a succession of $r + 1$ switch counts spaced uniformly by the interval i. The period covered will then be ri units long, and we shall assume that the call count has covered exactly the same total interval so as to eliminate certain of the correction difficulties described under IV-b above. Suppose the first switch count (at the beginning of the period) showed m calls up, the last switch count (at the exact end of the period) showed w calls up, and the sum of all of the $r + 1$ counts, *including* the first, totalled the number s.

As we found in equations (5) and (6) the average correction in the number s to be made on account of calls from the previous period being switch counted is

$$\bar{s}_m = \frac{m}{1 - e^{-\frac{i}{t}}},$$

and the standard deviation in the correction here is

$$\sigma_m = \sqrt{m}\,\frac{e^{-\frac{i}{2t}}}{1 - e^{-\frac{i}{t}}}.$$

At the end of the hour the corresponding correction was found in equations (10) and (11) to have an average of

$$\bar{s}_w = \frac{we^{-\frac{i}{t}}}{1 - e^{-\frac{i}{t}}},$$

and a standard deviation of

$$\sigma_w = \sqrt{w}\,\frac{e^{-\frac{i}{2t}}}{1 - e^{-\frac{i}{t}}}.$$

These are quite independent corrections if the observation period is several times the average holding time, the usual case. Then our best estimate of the number of switch counts we would have obtained if all (and only) those associated with the r calls originated in the period had been counted is,

$$s' = s - \bar{s}_m + \bar{s}_w\,, \tag{28}$$

and the standard deviation of s' is

$$\sigma_{s'} = \sqrt{\sigma_m^2 + \sigma_w^2}. \tag{29}$$

We now obtain immediately a preliminary figure for the average holding time of the n calls as

$$\overline{t'_1} = \frac{s'i}{n}, \tag{30}$$

and a standard deviation of this average by

$$\sigma_{\bar{t}'_1} = \frac{\sigma_{s'} i}{n}. \tag{31}$$

This last standard deviation comprehends the uncertainty in the holding time average caused by our inability to measure exactly the number of switch counts which properly should be associated with the n calls. We must now modify this measure of dispersion by the added fluctuation inherent in the method of switch counting at finite intervals. These variations were found to cause no change in the "expected" or most likely value of the average holding time (as shown by $\bar{x} = 0$ in equations (18), (22) and (26)), but the σ_x's of equations (19), (23) and (27) showed sizeable uncertainties which must be included. Since σ_x is for errors in the measurement of individual calls we obtain the standard error of the average as

$$\sigma_{\text{avg}} = \frac{\sigma_x}{\sqrt{n}}. \tag{32}$$

This error is built up on each call throughout the period and hence is practically independent of those errors arising at the ends of the hour. Their joint effects are additive so we obtain the estimates of the final parameters of the average holding time as

$$\overline{t'} = \overline{t'_1} + 0 = \frac{(s - \bar{s}_m + \bar{s}_w)i}{n}, \tag{33}$$

$$\sigma_{\bar{t}'} = \sqrt{\sigma_{\bar{t}'_1}^2 + \sigma_{\text{avg}}^2} = \frac{1}{n}\sqrt{(\sigma_m^2 + \sigma_w^2)i^2 + n\sigma_x^2}. \tag{34}$$

If the holding times are exponential these equations may be written as

$$\overline{t'} = \frac{i}{n}\left[s + \frac{we^{-\frac{i}{\bar{t}}} - m}{1 - e^{-\frac{i}{\bar{t}}}}\right], \tag{35}$$

$$\sigma_{\bar{t}'} = \frac{1}{n\left(1 - e^{-\frac{i}{\bar{t}}}\right)} \times \sqrt{\left((m + w)i^2 + n\bar{t}(2\bar{t} + i)\left(1 - e^{-\frac{i}{\bar{t}}}\right)\right)e^{-\frac{i}{\bar{t}}} - n\bar{t}(2\bar{t} - i)\left(1 - e^{-\frac{i}{\bar{t}}}\right)}. \tag{36}$$

It may be noted that the unknown average holding time $\bar{l}$ enters into both equations (35) and (36). They are not very sensitive to this value, however, and a first approximation obtained from $\bar{l} = \frac{(s - m)i}{n}$ will usually suffice. Further refinement may be obtained by recalculating using the new value $\bar{l}'$ found from equation (35). The form of the distribution represented by the parameters of equations (35) and (36) is not known of

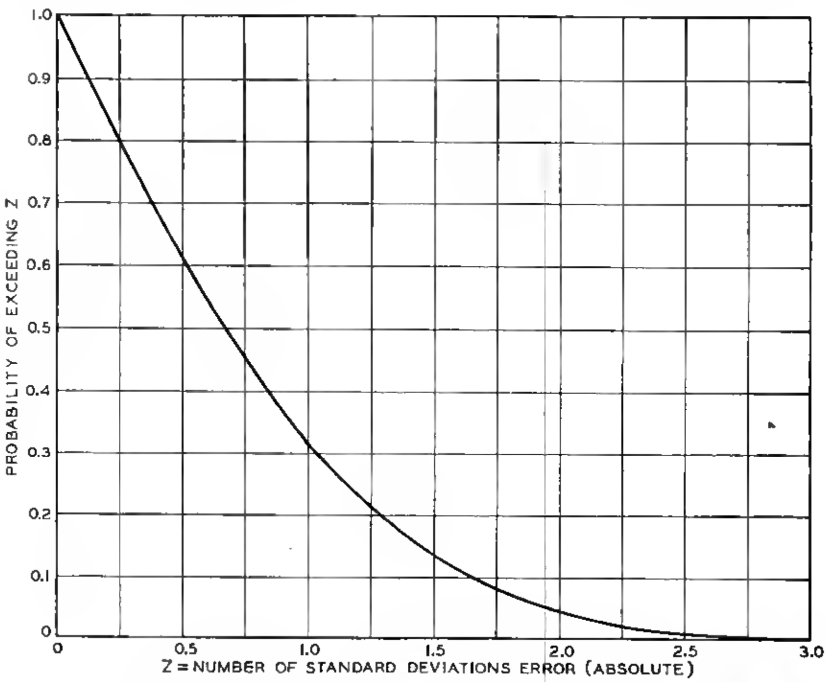

Fig. 16—Cumulative distribution of overall errors in average holding time, regardless of sign

course, but since the errors are essentially the sum of three primary error distributions which are inclined to be unimodal themselves (except for very small trunk groups), we shall probably be not far wrong to assume the normal form for $\bar{l}'$. If the magnitude only of the errors in the estimate $\bar{l}'$ of the unknown true holding time $\bar{l}$ of the n calls under observation is desired, we can readily construct a distribution of these discrepancies Figure 16 is the theoretical cumulative half-normal frequency curve, and when σ is equated to $\sigma_{\bar{l}'}$ found from equation (36), the probability of exceeding any given error in the holding time estimate may be read off directly.

Example: Suppose we have made 60-second switch counts on a group of

TA
SWITCH COUNT DAT

No.	At Beginning of the Hour			At End of the Hour						
						Due to calls which Pass #60 Sw. Count		Due to Calls Originating betwe #60 S.C. and End of the Hou		
	No. Calls Carried Forward m	Expected Switch Count Gain** $m\frac{1}{1-e^{-c}}$	St. Dev. of Total Gain** $\sqrt{m}\frac{e^{-\frac{c}{2}}}{1-e^{-c}}$	No. Calls Beyond #60 Count w	Distance between #60 and end of Hour (min.) j	Expected Loss** $w\frac{e^{-c}}{1-e^{-c}}$	St. Dev. of Total Loss** $\sqrt{w}\frac{e^{-\frac{c}{2}}}{1-e^{-c}}$	Avg. No. Expected $\alpha = \frac{j}{t}a_j$ $(a = 13.44)$	Expected S.C. Loss on these calls (Equation 12)	St. De of Tot S.C. L (Equat 13)
1	2	3	4	5	6	7	8	9	10	11
*	12	35.13	8.23	9	.4	17.34	7.13	2.24	4.71	4.7
2	19	55.61	10.35	24	.5	46.25	11.64	2.80	5.99	5.3
3	12	35.13	8.23	8	1.1	15.42	6.72	6.16	15.05	8.5
4	15	43.91	9.20	13	.9	25.05	8.56	5.04	11.78	7.5
5	10	29.27	7.51	13	1.0	25.05	8.56	5.60	13.40	8.07
6	18	52.69	10.08	9	1.0	17.34	7.13	5.60	13.40	8.0
7	14	40.98	8.89	10	1.0	19.27	7.51	5.60	13.40	8.0
8	13	38.05	8.56	7	.8	13.49	6.28	4.48	10.23	7.0
9	18	52.69	10.08	15	.7	28.91	9.20	3.92	8.78	6.5
0	15	43.91	9.20	7	.6	13.49	6.28	3.36	7.36	5.9
†*	19	55.61	10.35	24	.2	46.25	11.64	1.12	2.24	3.3
2	24	70.26	11.64	10	.5	19.27	7.51	2.80	5.99	5.3
3	13	38.05	8.56	19	1.0	36.61	10.35	5.60	13.40	8.0
4‡	19	55.61	10.35	10	.3	19.27	7.51	1.68	3.45	4.0
5	13	38.05	8.56	14	.5	26.98	8.89	2.80	5.99	5.3
6	15	43.91	9.20	12	1.1	23.13	8.23	6.16	15.05	8.5
7	21	61.47	10.88	11	.8	21.20	7.88	4.48	10.23	7.0
8	12	35.13	8.23	14	.75	26.98	8.89	4.20	9.50	6.7
9	13	38.05	8.56	8	.6	15.42	6.72	3.36	7.36	5.9
0	12	35.13	8.23	21	.5	40.47	10.88	2.80	5.99	5.3
1	14	40.98	8.89	16	.5	30.83	9.50	2.80	5.99	5.3
2	17	49.76	9.79	14	.94	26.98	8.89	5.26	12.43	7.7
3	12	35.13	8.23	13	.53	25.05	8.56	2.97	6.41	5.5
4‡	14	40.98	8.89	14	.8	26.98	8.89	4.48	10.23	7.0
5‡	18	52.69	10.08	9	.47	17.34	7.13	2.63	5.61	5.2
6*	18	52.69	10.08	6	.8	11.56	5.82	4.48	10.23	7.0
7	13	38.05	8.56	13	.53	25.05	8.56	2.97	6.41	5.5
8	14	40.98	8.89	16	.3	30.83	9.50	1.68	3.45	4.0
9	14	40.98	8.89	19	1.0	36.61	10.35	5.60	13.40	8.0
0	12	35.13	8.23	15	.4	28.91	9.20	2.24	4.71	4.7
1	15	43.91	9.20	16	.75	30.83	9.50	4.20	9.50	6.7

es: * In these hours additional errors are present because #1 switch count did not coincide well w
time 0.
† Hour No. 11 was only 59 minutes long.
‡ In each of these hours 1 pen registered call lasting for over 50 minutes was excluded as a trou
condition.
** c in the formulas at column headings stands for $\frac{i}{t}$, and was estimated as .418055 from $i = 60.2$ a
$t = 144$.

III

…on 31 Hours, Newark

…witch Counts in the Hour											
No. …witch …ounts …bs'd. s	Corr. No. s' $= s -$ (3) + (7) + (10)	St. Dev. of Corr. Total S.C. $[(4)^2+(8)^2+(11)^2]^{1/2}$	Length of S.C. Interval $i = 60j\,\dfrac{60-j}{59}$ (seconds)	No. Calls Originated in the Hour n	Estimated Average Holding Time $= \dfrac{s'\cdot i}{n}$ (seconds)	Std. Dev. of Avg. Holding Time due to Variation in No. Sw. Counts $\dfrac{(14)\cdot i}{n}$	Std. Dev. of Avg. Holding Time due to Errors in Measuring Each Call $= \dfrac{\sigma_x}{\sqrt{n}}$ (Equation 27)	Total Std. Dev. of Avg. Hold. Time $= [(18)^2 + (19)^2]^{1/2}$	Actual Measured Average Holding Time in the Hour (seconds)	Error in Estimated Holding Times in Multiples of the Theoretical Std. Dev. $\dfrac{(21)-(17)}{(20)}$	Perc… Err… (21) − (… (2…
12	13	14	15	16	17	18	19	20	21	22	2…
869	856	11.89	60.61	355	146.15	2.03	1.31	2.40	142.0	1.73	2.9…
837	834	16.48	60.51	359	140.57	2.78	1.30	3.05	139.9	.22	.4…
850	845	13.63	59.90	371	136.80	2.21	1.28	2.56	133.0	1.49	2.8…
848	841	14.67	60.10	328	154.10	2.69	1.36	3.02	147.2	2.28	4.…
837	846	13.96	60.00	311	163.22	2.69	1.40	3.04	159.9	1.09	2.…
868	846	14.74	60.00	372	136.45	2.38	1.28	2.71	131.8	1.72	3.5…
774	766	14.16	60.00	315	145.90	2.70	1.39	3.04	149.0	−1.02	2.…
842	828	12.75	60.20	332	150.14	2.31	1.35	2.68	143.7	2.40	4.4…
793	778	15.13	60.30	350	134.04	2.60	1.32	2.92	138.0	−1.36	2.8…
841	818	12.64	60.41	308	160.44	2.48	1.40	2.84	157.8	.93	1.…
865	858	15.92	60.81	343	152.11	2.82	1.33	3.09	149.0	1.01	2.…
901	856	14.86	60.51	374	138.49	2.40	1.28	2.73	131.1	2.71	5.…
811	823	15.67	60.00	341	144.81	2.76	1.34	3.07	144.0	.26	.…
801	766	13.43	60.71	318	146.24	2.56	1.38	2.88	144.5	.60	1.…
809	804	13.47	60.51	351	138.60	2.32	1.32	2.66	141.7	−1.16	2.…
802	796	15.01	59.90	305	156.33	2.95	1.41	3.28	158.7	− .72	1.4…
839	809	15.17	60.20	337	144.52	2.71	1.34	3.02	141.8	.90	1.…
792	793	13.88	60.25	332	143.91	2.52	1.35	2.86	141.7	.77	1.…
819	804	12.42	60.41	313	155.17	2.40	1.39	2.76	151.8	1.22	2.…
780	791	14.67	60.51	311	153.90	2.85	1.40	3.16	157.5	−1.14	2.…
805	801	14.08	60.51	380	127.55	2.24	1.26	2.56	128.9	− .53	1.…
863	853	15.34	60.06	303	169.08	3.04	1.42	3.36	160.5	2.55	5.…
734	730	13.12	60.48	290	152.24	2.74	1.45	3.08	158.0	−1.87	3.…
571	667	14.41	60.20	287	139.91	3.02	1.46	3.35	136.5	1.02	2.…
845	815	13.40	60.54	321	153.71	2.53	1.38	2.87	153.5	− .07	.1…
848	817	13.60	60.20	338	145.51	2.42	1.34	2.77	145.8	− .10	.…
805	798	13.33	60.48	324	148.96	2.49	1.37	2.83	150.5	− .55	1.…
766	759	13.64	60.71	316	145.82	2.62	1.39	2.94	150.0	−1.42	2.7…
755	764	15.85	60.00	340	134.82	2.80	1.34	3.11	137.6	− .90	2.…
783	781	13.23	60.61	337	140.46	2.38	1.34	2.72	1[illegible]5.2	−1.74	3.…
775	771	14.86	60.25	327	142.06	2.74	1.36	3.06	145.5	−1.13	2.3…

20 trunks for a period of one hour, finding $m = 7$ and $w = 13$ as the number of calls up at the beginning and the end of the hour, respectively. Suppose also we have a total of $s = 680$ switch counts, which includes the first count of 7 at time zero. If the register recording number of calls originated in the hour reads 282, what is the best estimate of the average holding time of the n calls, and what is the probability that the true holding time is witbin 3 seconds of this estimate?

We find our initial estimate for $\bar{t}$ from

$$\frac{(s - m)i}{n} = \frac{(673)60}{282} = 143.19 \text{ seconds.}$$

Substituting in (35) and (36) we find

$$\bar{t}' = 145.65 \text{ seconds,}$$

$$\sigma_{\bar{t}'} = 2.685 \text{ seconds.}$$

Then from Fig. 16 we read that the probability that this estimate of $\bar{t}$ is more than 3 seconds, that is $\frac{3}{\sigma_{\bar{t}'}} = 1.12$ standard deviations, in error is .263. Likewise we may read that the probability is .94 that the error in the average is *not* over 5 seconds.

As something of a final and overall comparison of theory and observation, the actual errors in holding times when estimated by the switcb count method for the 31 busy hours in Newark have been tabulated in Table III. The analysis of these pen register records was complicated by the fact that the intervals i varied somewhat from switch count to switcb count, and even more from hour to hour, so that tbe last switch count often came near the midpoint of the 60th minute. To some extent these irregularities of counting correspond more closely to the timing variations in manual switch counts than if they had been taken by machine at perfectly uniform intervals. Such corrections as could be managed by the application of equations (12) and (13) were made to the individual hours. In spite of these precautions the actual errors were somewhat larger than those which could be explained by theory although all large discrepancies were run down and accounted for. The absolute errors are shown plotted in Fig. 17a in terms of the theoretical standard deviations for each hour and in 17b in per cent of the observed holding times. This case will again serve to show that the switch count method is quite sensitive to variations from a perfect application of the rules, and that very considerable care is required to remain within the error limits indicated by the theory.

It is interesting to see what portion of the error is contributed by the two end effects and what by the errors made through "measuring" the calls by

switch counts at finite intervals. Hence on Fig. 18 is shown the overall error distribution for the illustrative example given above with the distributions of its two elements as well. The method of combining standard deviations (equation (34)) explains why the two error distributions of Fig.

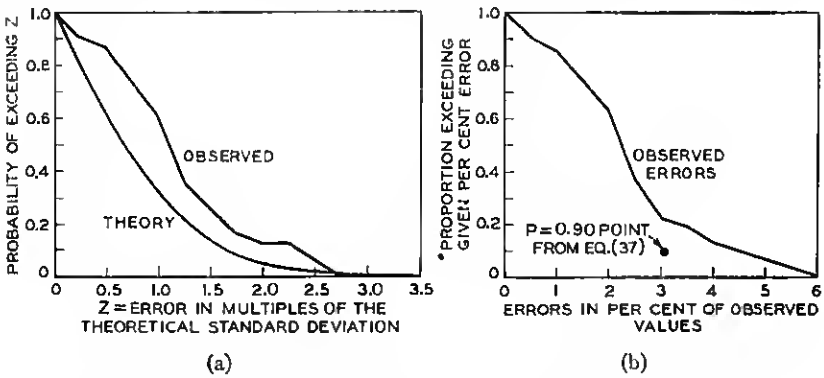

Fig. 17—Comparison of observed and theoretical overall errors

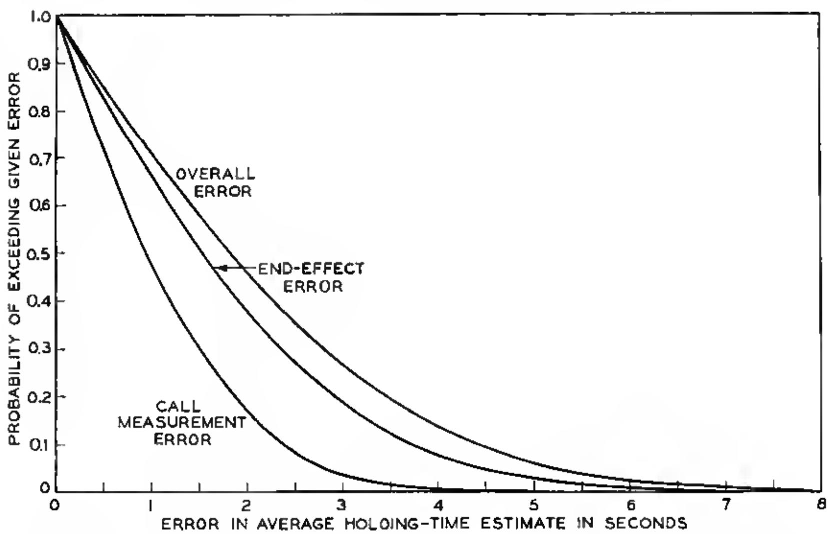

Fig. 18—Typical overall distribution of errors in estimating average holding times

18 do not add directly to give the overall error curve. We may immediately draw the important conclusion from Fig. 18 that the errors due to the uncertainties at the ends of the observation period may in certain cases be considerably greater than those due to the lack of precise measurement of the length of each individual call.

The relative prominence of these two errors will not change very rapidly with different sizes of observation periods (that is, the number of calls, n), since the end effects' standard deviation (equation 31) will vary inversely as n, while the standard deviation due to the error in measuring individual calls (equation 32) will vary inversely as the $\sqrt{n}$. Doubling the length of the observation period would then decrease the first σ to one-half, and the second σ to .707 of its former value.

Equation (36) shows that

$$\sigma_{\bar{l}'} = f(m, w, n, \bar{l}, i).$$

We will not know m and w before making the switch counts but we can probably substitute the average value of their sum which is $2n\bar{l}/T$, without seriously disturbing the average value of $\sigma_{\bar{l}'}$. This gives

$$\sigma_{\bar{l}'} \approx \frac{\sqrt{\frac{\bar{l}}{n}}}{1 - e^{-\frac{i}{\bar{l}}}} \sqrt{\frac{2\bar{l}^2 e^{-\frac{i}{\bar{l}}}}{T} + \left[(2\bar{l} + i)e^{-\frac{i}{\bar{l}}} - 2\bar{l} + i\right]\left(1 - e^{-\frac{i}{\bar{l}}}\right)}, \quad (37)$$

where now $\sigma_{\bar{l}'}$ is a function of only four variables, n, $\bar{l}$, i and T.

It is of interest to compare the errors predicted by equation (37) with those found by the theory formulated for the particular m and w observations found in each hour's observations at Newark. An average of $n = 332$ calls per hour was observed with an average holding time in the order of $\bar{l} = 145$ seconds. The switch count interval, i, was approximately 60 seconds, and the observation period was $T = 3600$ seconds. If we take $P = .90$, the per cent error corresponding will be $\pm 100\ (1.645)\ \sigma_{\bar{l}'}/\bar{l}$. Using the above estimate of $\sigma_{\bar{l}}$, we find an error of about 3.05 per cent, or $\pm .0305$ $(145) = \pm 4.42$ seconds. This point has been plotted on Fig. 17b and bears about the same relationship to the observed errors as does the theoretical curve in Fig. 17a in which the comparison takes into account the actual calls carried beyond the start and end of each hour. The discrepancy in Fig. 17b is largely accounted for by the same discussion given heretofore for Fig. 17a.

The Overall Error in Estimating the Average Holding Time

The engineer who has the problem of devising a switch count schedule will want to be able to estimate at least roughly the order of accuracy he will actually obtain in the average holding time found from the data in a number of observation periods. Up to this point in this section we have concerned ourselves with discovering only the errors inherent in measuring the average length of a *particular n calls* of the exponential type in an observation period of length T. As we saw in section III, even when such an

average length $\bar{l}$ is accurately known for the n calls, it may not exactly or even closely coincide with the true average of the universe of calls of wbich the n are presumed to form a random sample. The errors we have just studied and those described in section III must now be combined to give us a measure of the overall accuracy of the switch count method.

Equation (2), when applied to the exponential holding times we are here concerned with, gives us for the standard error of the average in a sample

$$\sigma_{\text{sampling}} = \frac{\sigma}{\sqrt{n}} \approx \frac{\bar{l}}{\sqrt{n}}. \tag{38}$$

This error is independent of that represented by $\sigma_{\bar{l}'}$, so we may determine the overall (oa) standard error by

$$\sigma_{oa} = \sqrt{\sigma^2_{\text{sampling}} + \sigma_{\bar{l}'}{}^2}. \tag{39}$$

We shall be particularly interested in knowing how much the value of $\sigma_{\bar{l}'}$ contributes to the overall standard deviation, σ_{oa}. This may be conveniently expressed by writing tbe ratio

$$q = \frac{\sigma_{oa}}{\sigma_{\text{sampling}}} = \sqrt{1 + \frac{\sigma_{\bar{l}'}{}^2}{\sigma^2_{\text{sampling}}}}$$

$$= \sqrt{1 + \frac{\frac{2i}{\bar{l}}\frac{i}{T}e^{-\frac{i}{\bar{l}}} + \left[\left(2 + \frac{i}{\bar{l}}\right)e^{-\frac{i}{\bar{l}}} - 2 + \frac{i}{\bar{l}}\right]\left(1 - e^{-\frac{i}{\bar{l}}}\right)}{\left(1 - e^{-\frac{i}{\bar{l}}}\right)^2}}. \tag{40}$$

Now it is readily seen from (40) that q depends on $\bar{l}$, i and T. Hence if T is held constant we may plot curves between q, i and $\bar{l}$ as shown on Fig. 19. What is more, if T is varied, say increased by a factor k, equation (40) shows that if i and $\bar{l}$ are also increased by the same factor the values of q resulting may still be read directly from Fig. 19.

For example: If approximately 100-second exponential calls are to be switch counted for an hour with observation intervals of 120 seconds, we read on Fig. 19 that the *overall* standard error (or P = .50, .90, .99, etc. error) in estimating the true average holding time is q = 1.134 times the basic standard error resulting from taking a random sample of n calls from a very large universe of calls. That is to say, the residual sampling error present in a stop watch measurement of the n calls is increased by 13.4 per cent due to our resort to switch count methods.

If a continuous period of T = 2 hours (i.e. k = 2) is switch counted in just the same manner and under the same conditions, we should now read on the $\bar{l}$ = 50 seconds curve opposite i = 60 seconds, giving us an increase over the basic sampling error of 12.4 per cent. This meets one's common

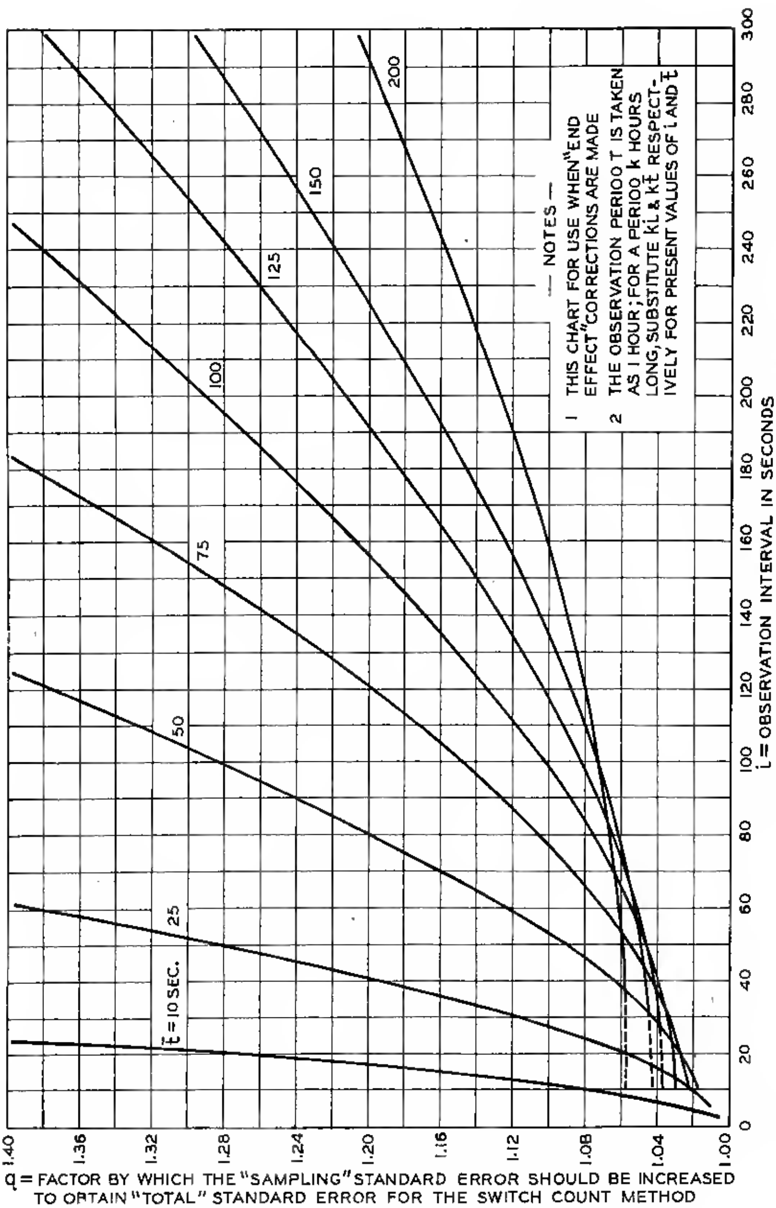

Fig. 19—Increase in "overall error" over "sampling error" in the estimate of holding time averages by switch counts (with "end effect" corrections)

sense conclusions since, as previously suggested, when the observation period is increased by a factor k,

a. The "sampling" error decreases as $\frac{1}{\sqrt{k}}$, since we are looking at k times as many calls,

b. The error of measurement on the individual calls decreases also as $1/\sqrt{k}$, and for the same reason, and

c. The "end effect" errors are unchanged in actual magnitude but are now prorated over k times as many calls; hence, the effect of this element decreases as $1/k$.

The overall effect then as k increases is for σ_{ou} to decrease faster than σ_{sampling} as we have just seen in the example. Thus not only is the sampling error decreased by lengthening the observation period, but the overall error decreases even faster. It is clear then that an observation period as long as is consistent with a "controlled" universe of holding times is to be preferred.

Further important deductions may be drawn from a study of Fig. 19. If we wish to minimize the effect of errors introduced by a use of the switch count method we shall need to select our observation interval i so that it will be relatively small compared with the average holding time. Apparently we should do well to keep $\frac{l}{i} \geqq 1.5$; the higher this ratio the better, although the improvement beyond, say, 2.0, is slow. Roughly, for commonly observed local subscriber holding times if the holding time is at least twice the observation interval the increase in error occasioned by the switch count method over the stop watch method need not exceed 7 per cent.

Fig. 19 has been constructed for use when l is estimated as l' from equation (35),

$$l' = \frac{(s - s_m + s_w)i}{n} = \frac{i}{n}\left[s + \frac{we^{-\frac{i}{l}} - m}{1 - e^{-\frac{i}{l}}}\right], \tag{35}$$

in which s is the sum of $r + 1$ switch counts (which includes counts at both the extreme ends of the period), m is the count at the beginning, and w the count at the end of the period.

It will be of interest to estimate something of the enlarged error when l' is found merely by taking

$$l' = \frac{(s - m)i}{n} \tag{41}$$

which entirely neglects the special information contained in the first and last switch counts. $\left(\text{If only } r\left(= \frac{T}{i}\right) \text{ counts are made they should be at the}\right.$

end of each i-interval, the last one coming at the exact end of the whole period.) This is the common case in which we merely sum all the switch counts, multiply by the counting interval and divide by the number of calls shown on the peg count meter as originating in the period T. The standard error for each end effect will then be approximately that given by σ'_m in equation (6′) where no attention is paid to the end switch count values of m and w. Substituting σ'_m for σ_m and σ_w in (29) and following the same analysis as before gives for q' (instead of q),

$$q' = \sqrt{1 + \frac{\frac{2i}{\bar{l}}\frac{i}{T}\left(e^{-\frac{i}{\bar{l}}} + 1\right) + \left[\left(2 + \frac{i}{\bar{l}}\right)e^{-\frac{i}{\bar{l}}} - 2 + \frac{i}{\bar{l}}\right]\left(1 - e^{-\frac{i}{\bar{l}}}\right)}{\left(1 - e^{-\frac{i}{\bar{l}}}\right)^2}}. \qquad (42)$$

A plot of this last expression is given in Fig. 20. By comparing points on Fig. 20 with corresponding ones on Fig. 19, one obtains an idea of the increase of error due to failure to correct the switch counts for the end effects as indicated in equation (35). For example, with 100-second calls switch counted at 120-second intervals we find $q = 1.134$ while $q' = 1.203$, indicating quite a marked increase in the overall error. The particular errors resulting in any given circumstance coupled with the cost of making the end effect corrections will determine the practical desirability of which method to adopt, that is whether the factor for increasing the basic sampling standard error shall be read from Fig. 19 or from Fig. 20.

Finally, a chart has been drawn up as Fig. 21, by which a measure of the overall error in estimating the unknown true holding time may readily be determined. The right hand section of the chart is a redrawing of Fig. 5 given in section III for the sampling errors of individually measured calls. Scale A is carried across and reproduced at C permitting the small nomograph B C D to give easily the product of the sampling error and the q (or q') factor at D. The left hand chart is based simply on the fact that the overall error decreases inversely as the square root of the number of periods switch counted. From it the number of periods required to obtain any desired accuracy can be read.

The estimate of the average holding time will be found from a simple average of the estimates made for individual observation periods,

$$\bar{l}' = \frac{\bar{l}'_1 + \bar{l}'_2 + \cdots + \bar{l}'_g}{g}. \qquad (43)$$

If a certain per cent error in the estimated average holding time is obtained for a single period the improvement for the combination of g periods

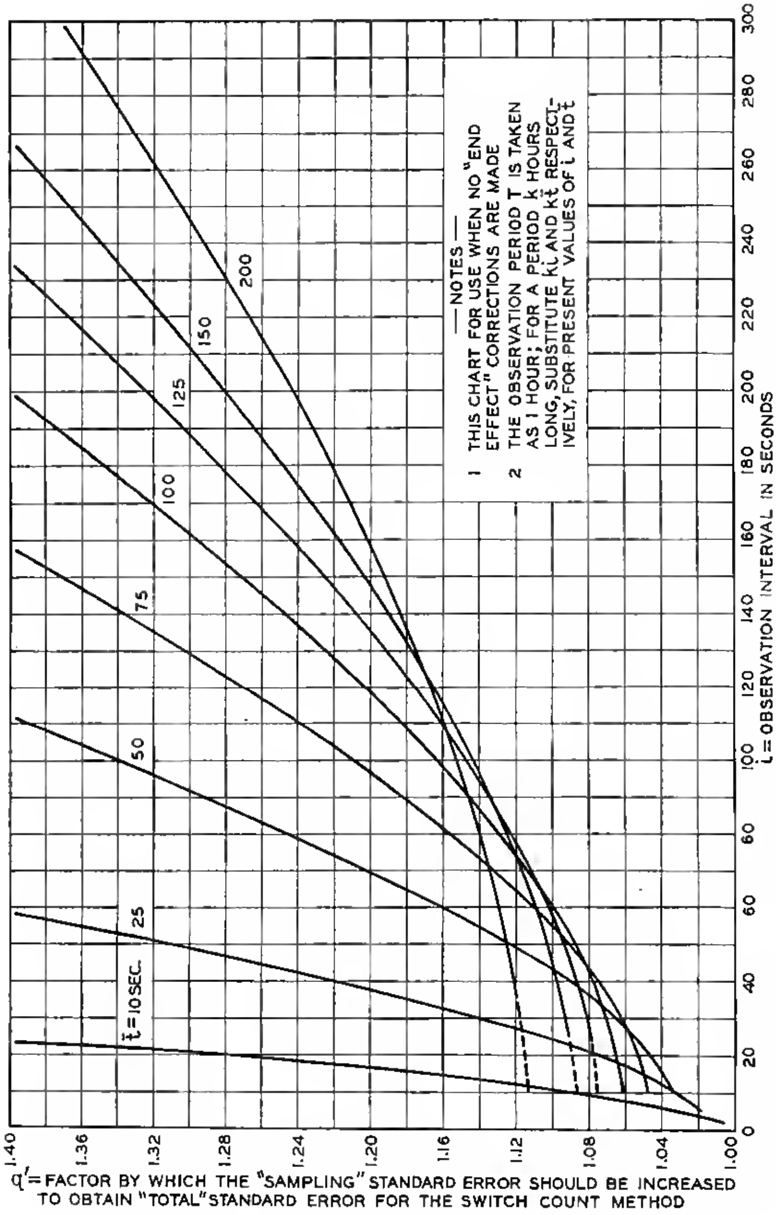

Fig. 20—Increase in "overall error" over "sampling error" in the estimate of holding time averages by switch counts (no "end effect" corrections)

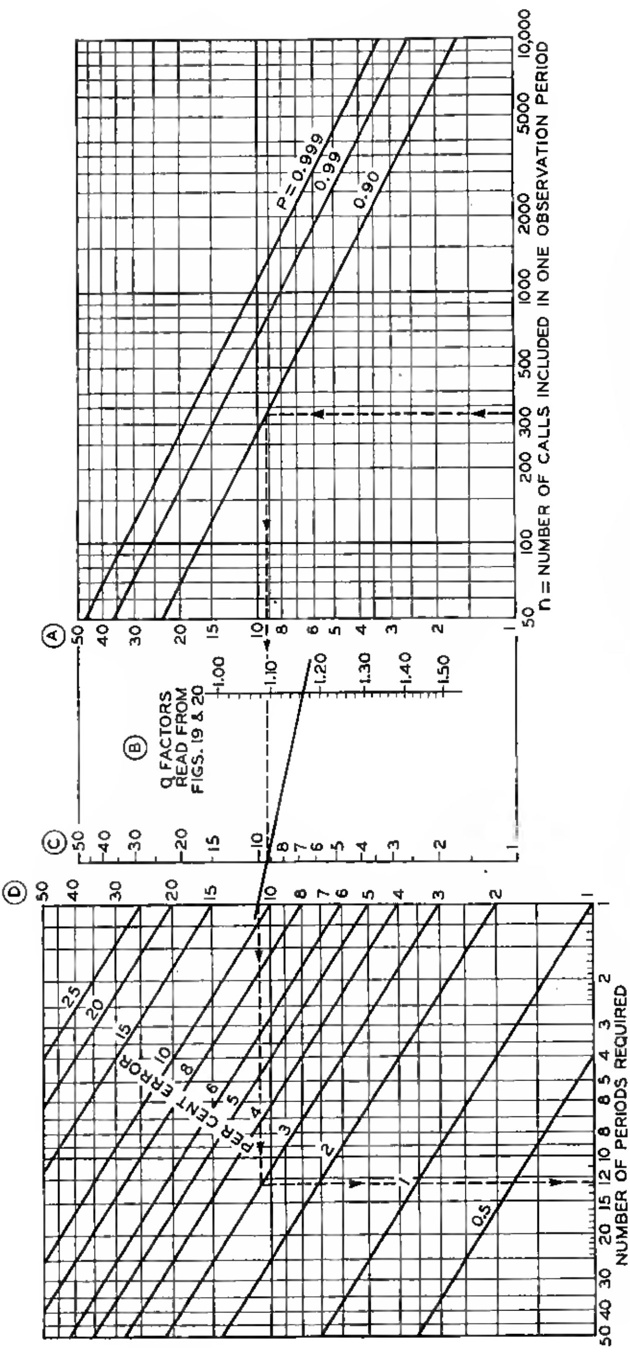

Fig. 21—Determination of number of observation periods required to produce any desired accuracy in estimating holding times by switch counts

is then,

$$\frac{1}{\sqrt{g}} \text{ (single period error).}$$

Illustrative Example. If calls of approximately 135 seconds holding time are to be switch counted at 3 minute intervals by 1 hour periods, how many such periods will be required to give us an assurance of $P = .90$ that the resulting holding time estimate does not differ from the true by more than 3 per cent? Assume an average of 325 calls per hour over the group, and that end effect corrections (a) will be made, and (b) will not be made.

Solution (a) We first read on Fig. 19 that $q = 1.165$. Turning to Fig. 21, we find that opposite $n = 325$ and $P = .90$ we have an error of 9.2 per cent, which we carry over to the C scale of the nomograph. Laying a straight-edge across this point and $q = 1.165$ determines a point on D somewhat above the 10 per cent line. Projecting this point across to the desired 3 per cent error line and dropping to the lower edge of the chart we find that 12.7, say 13, such one-hour observation periods will be necessary to meet the accuracy specifications of the problem. This is the effort required when the end effect corrections are made.

(b) If the end effect corrections are to be ignored, we determine our factor from Fig. 20 to be $q' = 1.268$. Proceeding on Fig. 21 exactly as before we find that the number of one-hour observation periods required is increased to 15.1, say 15. Thus a failure to make the end effect corrections causes us to increase our switch counts by about 20 per cent. It is by such comparisons as these that one decides the practical desirability of making the end effect corrections.

Summary of Switch Count Procedures for Exponential Holding Times

The choice of the size of unit observation period should rest primarily on the considerations discussed in Section II, that is the homogeneity of the holding time data from hour to hour, day to day, etc. Other things being equal, we should select the longest period consistent with the view that enough periods must be included so that representative sampling of all known or suspected major variations in holding time character is accomplished. The length of the switch count interval will likely already have been decided by the equipment at hand or by other considerations. If a choice is available, however, a short interval will produce more reliable results than a long one, by the amounts indicated on Figs. 19 and 20. With these matters decided, Fig. 21 is consulted to see how many periods must be switch counted in order to obtain the desired accuracy.

Having actually made the switch counts, exercising the cautions we have mentioned, the average holding time $\bar{t}'$ for each period is obtained from equation (35) if the correction of the counts for the end effects in each

period is made. If no such corrections are to be made, equation (41) is used instead of (35). The arithmetic mean, equation (43), of the values obtained in the various periods will then be the best available estimate of the unknown true average holding time. The reliability of this figure should be substantially that which the schedule was designed to produce.

Switch Count Errors for Non-Exponential Holding Times

If switch counts are made on calls with other than exponential holding times the resultant errors may be greater or less than those shown by Figs. 19, 20 and 21. The comparison of typical σ_x's calculated from equations (23) and (27) would suggest that for varying holding times the error in the measurement of individual calls is not greatly dependent on the form of the holding time distribution *as long as the average call length covers several intervals i*. In such a case the charts developed for exponential holding times can probably be used with little allowance for the discrepancy present.

On the other hand for calls with an unusual or extreme fluctuation about an average $\bar{t}$ less than i, the errors due to assuming the situation to be equivalent to the exponential case may be no longer negligible. The only procedure then would appear to be either to work out the errors actually present, reverting to the basic error equations (14) and (15), and approximating the new end effect corrections, or to revise the switch counting program to materially shorten the interval i.

For relatively constant holding times the value of σ_x can be reduced to a small figure by choosing the switch count interval i so that it is contained in the average holding time $\bar{t}$ closely a whole number of times. Then by equation (19) the error in individual call measurements must, of necessity, be small in nearly every instance. It will be noted that the above specification permits choosing $i = \bar{t}$; moreover, it may readily be seen that in this case the end effect corrections will tend to disappear, giving a highly accurate measurement with relatively few observations. Just how many, of course, will depend on how constant is the quantity measured, and how closely the switch count interval i approaches the true average $\bar{t}$.

V—General Summary

The general problem of determining the average holding times of subscribers' or other calls by sampling methods has been discussed. The need for a proper apportionment of the sample is emphasized and examples are given from telephone experience to illustrate typical analysis procedures. Methods for estimating the reliability of these sampling results for both directly measured holding times and for switch count studies are given along with various curves and charts calculated to assist the traffic engineer in devising a working schedule for the sampling of holding times, particularly those of an exponential character.